

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: December 6, 2002, 19:53:16 : Search time 2690 Seconds
(without alignments)
16498.803 Million cell updates/sec
Title: US-10-025-514-15
Perfect score: 1525
Sequence: 1 tctagaccatgaagaccct.....ccagtcaaggcctagtcgac 1525

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl : *
1: gb_ba : *
2: gb_htg : *
3: gb_in : *
4: gb_om : *
5: gb_ov : *
6: gb_pat : *
7: gb_ph : *
8: gb_pl : *
9: gb_pr : *
10: gb_ro : *
11: gb_sts : *
12: gb_sy : *
13: gb_un : *
14: gb_vl : *
15: em_ba : *
16: em_fun : *
17: em_hum : *
18: em_in : *
19: em_mu : *
20: em_om : *
21: em_or : *
22: em_ov : *
23: em_pat : *
24: em_ph : *
25: em_pl : *
26: em_ro : *
27: em_sts : *
28: em_un : *
29: em_vl : *
30: em_htg_hum : *
31: em_htg_inv : *
32: em_htg_other : *
33: em_htg_mus : *
34: em_htg_pln : *
35: em_htg_fod : *
36: em_htg_mam : *
37: em_htg_vrt : *
38: em_sv : *
39: em_hgo_hum : *
40: em_hgo_mus : *
41: em_hgo_other : *

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match %	Length	DB ID	Description
1	629.4	41.3	1308	6	AR111412 Sequence
2	628.4	41.2	1185	6	AR111411 Sequence
3	437	28.7	1345	9	HUMALATB
4	433.8	28.4	1434	6	E00631 DNA encodin
5	432.2	28.3	1312	6	I02706 Sequence 1
6	430.6	28.2	1584	9	BC011991 Homo sapi
7	429	28.1	1185	6	AR111410 Sequence
8	429	28.1	1352	6	AX335339 Sequence
9	429	28.1	1352	9	HUMALATM
10	429	28.1	1371	6	AX335338 Sequence
11	429	28.1	1371	9	HSATPR1
12	429	28.1	1399	9	AK026174 Homo sapi
13	429	28.1	1434	6	E00195
14	429	28.1	1434	6	I04196
15	429	28.1	1434	6	I04272 Sequence 1
16	429	28.1	1434	6	I07849 Sequence 2
17	429	28.1	1435	6	AX019486 Sequence
18	429	28.1	2478	17	AF130088 Homo sapi
19	427.4	28.0	1337	9	HUMALAT2
20	427.4	28.0	1378	6	I02398 Sequence 8
21	427.4	28.0	1431	9	BC015642 Homo sapi
22	427.4	28.0	2571	17	AF113676 Homo sapi
23	425.8	27.9	1299	6	I07949 Sequence 2
24	425.8	27.9	1308	9	BABAIATA
25	425.8	27.9	1317	6	I00556 Sequence 2
26	425.8	27.9	1378	6	I03509 Sequence 6
27	425.8	27.9	1378	6	I07947 Sequence 4
28	425.8	27.9	1434	6	I01352 Sequence 1
29	425.4	27.9	1185	6	A01846 Human mRNA
30	424.2	27.8	1378	6	I01227 Sequence 2
31	417	27.3	1191	9	AB004044 Cercopith
32	407	26.7	1390	6	AX202089 Sequence
33	404.6	26.5	1356	6	I36163 Sequence 16
34	403	26.4	1356	6	I36161 Sequence 12
35	402.6	26.4	2013	6	AX472008 Sequence 18
36	399.8	26.2	1356	6	I36164 Sequence 18
37	390.2	25.6	1351	10	AB000550 Spermo
38	387	25.4	1372	10	AB000552 Callosciu
39	385.4	25.3	1351	4	SSANTIOLA
40	378	24.8	1306	10	RATATRA1
41	377.2	24.7	1352	10	S77822
42	376.4	24.7	1380	10	RATAPI
43	372.8	24.4	1343	10	AB000546
44	372.6	24.4	1242	6	E13268
45	370.6	24.3	1380	4	BTALAT

ALIGNMENTS

RESULT 1
AR111412
LOCUS AR111412 1308 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 6 from patent US 6127145.
ACCESSION AR111412
VERSION AR111412.1 GI:12828260
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1308)
AUTHORS Sutliff,T.D. and Rodriguez,R.L.
TITLE Production of .alpha..sub.1 -antitrypsin in plants
JOURNAL Patent: US 6127145-A 6 03-OCT-2000;
FEATURES Location/Qualifiers

Mon Dec 9 12:50:38 2002

[illegible]

RESULT 2	AR111411	AR111411	1185 bp	DNA	linear	PAT 14-FEB-2001
LOCUS	Sequence 3 from patent US 6127145.					
DEFINITION	AR111411					
ACCESSION	AR111411					
VERSION	AR111411.1	GI:12828259				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 1185)					
AUTHORS	Sutliff,T.D. and Rodriguez,R.L.					
TITLE	Production of alpha. sub.1 -antitrypsin in plants					
JOURNAL	Patent: US 6127145-A 3 03-OCT-2000;					
FEATURES	location/Qualifiers					
source	1. .1185					
BASE COUNT	276 a	396 c	334 g	179 t		
ORIGIN	/organism="unknown"					
Query Match	41.2%; Score 628.4; DB 6; Length 1185;					
Best Local Similarity	70.7%; Pred. No. 7.9e-124; Indels 0; Gaps 0;					
Matches 836; Conservative						
QY 12	GAAGACCTCTAAGGCGAGCGCGCTCAAAAAACCGACACGAGTCATCAGCACCAAGACCAT 71					
Db 1	GAGGACCCGAGGCGAGCGCCGCCAGAGACCGACCCAGCAGCACCAGCAGCAGGACAC 60					
QY 72	CGACTCTTAAATAAATTAATCTCCAAATTTAGCGGAATTTGCTTTCTTTGTATAGACAA 131					
Db 61	CGAGCTTCAACAAGATCACCOCGAATTTGGCGAATTCGCCTTCAGCTGTACCGCCAG 120					
QY 132	TTAGCTCATCAAGTAATTTCTACTAACATTTTTTTTAGTCCTGTGTTCTATATGCCACTGCT 191					
Db 121	CTCGGCGACCATTCGCACTCCACACACATCTTCTTCAGCCCGTGAGCATCGCCACCGCC 180					
QY 192	TTTCGCCATGTTGAGTTTAGTGTACTAAAGCCGATACCCATGACGAGATTTTAGAAGTTTA 251					
Db 181	TTTCGCCATGCTGCTCCCTGGTACCAAGGCGGACACCCACGACGAGATCTCTCGAAGGCGTG 240					
QY 252	AACTTTAATTTGACCGAAATCCGAGAGCCCAAAATTCAGAGGGTTTTCAAGAGTTGTTG 311					
Db 241	AACTTCAACTGACGAGATCCCGGAGGCGAGATCCAGGAGGGCTTCGAGGAGTCTCTC 300					
QY 312	AGAACTTTCAATCAACTGATTTCTCAATTCGAATTAACCTACTGGTAACGGTTTATTTTG 371					
Db 301	AGAGCGCTCAACGACGCGGACTCCAGCTCCAGCTCACCAACCGCAACGGCTCTTCCCTG 360					
QY 372	TCTGAAGGTTTAAATTTGGTTGACAAATTCCTAGAGAGCGTCAAGAACTATATCATAGT 431					
Db 361	TCGAGGGCTCAAGCTCGTCGATAAGTTTCTCGAGGAGCTGAAGAGAGTCTACCACTCC 420					
QY 432	GAGGCTTTTACCGTTTAATTTTGGTGATCTAGGAGAGCTAAAAAGCAAAATTAATGATTAT 491					
Db 421	GAGGCGTTCAACCTCAACTTCGGGGACACCGAGGCGCCAAAGAGCAGATCAACGACTAC 480					

QY	492	GTGAGAAAGGACCCAGGCTAGTTCGTTAGACCTAGTTAAGAAATATAGATCGTGATACC	551
Db	481	GTGAGAAAGGACCCAGGCTAGTTCGTTAGACCTAGTTAAGAAATATAGATCGTGATACC	540
QY	552	GTCTTCGCACTAGTTAACTATATATTTTTCAGAGGCTAAGTGGGAACGTCCTTTCAGGTT	611
Db	541	GTCTTCGCGCTCGTCAACTACATCTTCTTCAAGGCAAGTGGGACGCCGCTTCAGGTT	600
QY	612	AAAGATCTAGAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCCTCAAGTTCCAATG	671
Db	601	AAAGATCTAGAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCCTCAAGTTCCAATG	660
QY	672	ATGAAAGACTGGGTATGTTCAATATTCACCAATTCGCAAAATTAAGTTCTTGGGTCTTA	731
Db	661	ATGAAAGACTGGGTATGTTCAATATTCACCAATTCGCAAAATTAAGTTCTTGGGTCTTA	720
QY	732	TAAATGAAGTATTTAGGTAACGCTACTGCTATTTTTCACCAAGGTAAGGTAAGCTT	791
Db	721	CTCATGAAGTATGTTGGGAACGCCACCGCATCTTCTTCTGCGGACGAGGCAAGCTC	780
QY	792	CAACATTTAGAGAACTAGTTCGCTCATGACATATTAATTAATTTTATAGAAACGAGGAT	851
Db	781	CAGCACTGGAGAACGAGTTCGCTCATGACATATTAATTAATTTTATAGAAACGAGGAT	840
QY	852	CGTCTAGGCTTCTCTGACCTGCGCAAGTTAAGTATCACCGTACTTAGCACTTAAAA	911
Db	841	AGGCGCTCGCTAGCTTCACTTCCGAGCTGACATCACCGGACGCTAGACCTGAAG	900
QY	912	TCTGTTTATAGGCGAGTATTAACCAAGTTTTCCTAAGGTCGCGATTTAGTGGT	971
Db	901	AGCGTCTGGGCGAGTTCGCTATCAAGAGTCTTTCAGCAACGCGGCGGACCTTCCCGG	960
QY	972	GTACTGAGAAAGTCCATTAATTAAGTATTAAGTATTCACCAAGGCTGTCACCAAGCGCTT	1031
Db	961	GTAGCGGAGGCGGCGGCTGAGCTCTCAAGGCGGCTGACCAAGGCGGCTGTCACGATC	1020
QY	1032	GATGAAAGGCTACCGAGGCGGCGGCTGATGTTCTTGGAGCTTATTCCAATGAGCAATT	1091
Db	1021	GACGAGAGGCGGAGGCTGCGGCGGCTGATGTTCTTGGAGGCTATTCCTGATGCTTCCATC	1080
QY	1092	CCACGAGAGTTAAATTAATTAAGTATTCGTTTTCGATGATGAGCAAGCACTAAA	1151
Db	1081	CGGCGGAGGCTCAAGTTCACCAAGGCTTCTGCTTCTGATGATGAGCAAGCAAGG	1140
QY	1152	AGCCCATTTTATGGTAAAGTTGTCACCAAGGCTGATGATGATGATGATGATGATGATG	1193
Db	1141	AGCCCATTTTATGGTAAAGTTGTCACCAAGGCTGATGATGATGATGATGATGATGATG	1182
RESULT 3			
HUMALATB			
LOCUS	HUMALATB 1345 bp mRNA linear PRI 30-OCT-1994		
DEFINITION	Human alpha-1-antitrypsin mRNA, complete cds.		
ACCESSION	M11465		
VERSION	M11465.1 GI:1177826		
KEYWORDS	alpha-1-antitrypsin.		
SOURCE	Human liver, cDNA to mRNA.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	Cell-specific expression of a transfected human alpha 1-antitrypsin		
JOURNAL	Cell 41 (2), 531-540 (1985)		
MEDLINE	85176977		
PUBMED	2985281		
REFERENCE	2 (sites)		
AUTHORS	Curjel,D., Brantly,M., Curiel,E., Stier,L. and Crystal,R.G.		
TITLE	Alpha 1-antitrypsin deficiency caused by the alpha 1-antitrypsin		
JOURNAL	Nullmattawa gene. An insertion mutation rendering the alpha		
	1-antitrypsin gene incapable of producing alpha 1-antitrypsin		
	J. Clin. Invest. 83 (4), 1144-1152 (1989)		

MEDLINE

PUBMED

COMMENT

FEATURES

source

gene

CDS

variation

variation

variation

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

[2] sites; alpha-1-antitrypsin deficiency disease mutations. The Granite Falls (a one base pair deletion), Bellingham (a substitution) and Mattawa (a one base pair insertion) mutations all result in premature transcription termination.

Location/Qualifiers

1. .1345

/organism="Homo sapiens"

/db_xref="taxon:9606"

/map="14q32.1"

1. .1345

/gene="PI"

15. .1268

/gene="PI"

/note="alpha-1-antitrypsin"

/codon_start=1

/protein_id="AA51546.1"

/db_xref="GI:177827"

/db_xref="GDB:G00-120-289"

/translation="MPSSVSGILLAGLCLVPSVLAEDPQGDAAOKTDTSHDDHP
TFNKITPNAEFAFSLYRLAHQSNSTIFFSPVSIATAFAMLSLGTAKDTHEILEG
LNFNLTEIPAOIHEGQELLRTLNQDSQLQTLTGNGFLSEGLKLVKDFLEVKKL
YHSEAFVNFQDHEEAKKOINDYVEKGTQKIVDLVKELDRDTFALVNY IFEKQWE
RPEVKDTEDEDFHVDVTVKVPKMKRLGMENIOHCKKLSKVLMLKYLGNATIFF
LPDEGLQHLNELHLDITIKFLENDRRSASLHLPKLSITGYDLSKVLGQIGITKV
FVNGADLSQVTEAPUKLSKAVHKAVLTIDKGTGAAGAMFLEAIPMSIPPEVKFNKP
FVPLMTQNTKSPLEFMKVNPQK"

562. .564

/gene="PI"

/note="acg in wt; ag in null-Granite Falls"

732

/gene="PI"

/note="a in wt; t in null-Bellingham"

1141

/gene="PI"

/note="ta in wt; tta in null-Mattawa"

354 a 381 c 317 g 293 t

28.7%; Score 437; DB 9; Length 1345;

60.6%; Pred. No. 5.2e-83;

Conservative 0; Mismatches 465; Indels 0; Gaps 0;

12 GAAGACCTCAAGCGGACGCGCTCAAAAAACCGACACAGTATCATCAGCAGCAGACCAT 71

84 GAGGATCCCCAGGAGATGCTGCCAGACAGATACATCCACCATGATCAGGATCAC 143

72 CCGACTTTTAAATTAATTTACTCCAAATTTAGCCGAATTTGCTTTTCTTTGTATAGACAA 131

144 CCAACCTTCAACAGATACACCCCAACCTGGCTGAGTTCGCTTACGCCCTATACCCCGAG 203

132 TTAGCTCATCAAAAGTAAATTTCTACTAATTTTCTTTAGTCTCTTCTTCTTATGCACTGCT 191

204 CTGGCAGCAGTCCACACACCATATCTTCTTCTCCCATGAGCAGTCTGATCAGCC 263

192 TTGCGCATGTTAGTTTGTAGTACTTAAAGCGGATACCATGACGAGATTTTGAAGTTTA 251

264 TTTGCAATGCTCTCCCTGGGAGCAAGGCTGACATCAGCATGAAATCTCTGGAGGCGCTG 323

252 AACTTTAATTTGACCGAATTTCCAGAGGCGCAATTTACAGNGGGTTTTCAGAGAGTTGTTG 311

324 AATTTCAACCTCAGGAGATTCGCGAGGCTCAGATCCATGAAGGCTTCCAGGAACCTCTC 383

312 AGAATTTGAATCAACCTGATTTCTCAATTTAACTACTGTTAACTGTTTATTTTGTG 371

384 CGTACCTTAACAGCAGCAGACACCATCCAGCTGACCCCGGCAATGGCTGTTCTCTC 443

372 TCTGAAGGTTTAAATTTGGTTTGAATAATTCCTAGAGAGCGTCAAGAACTATATCATAGT 431

444 AGCGAGGCGCTGAAGCTAGTGGATAGTTTGGAGGATGTTTAAAAAGTTGTACCATCA 503

432 GAGGCTTTTACCCTTAATTTTGGTGATCTGAGGAAGCTTAAAAAGCAAAATTAATGATTAT 491

Db 599 GTGGAGAGGGTACTCAAGGGAAATTTGGGATTTGGTCAAGGAGCTTGACAGAGACACA 658
QY 552 GTCTTCGGCACTAGTTAACTATATTTTTCAGGGTAAAGTGGGAAGCTTTCAGAGTT 611
Db 659 GTTTTTCCTGCTGGTGAATTTACATCTCTTTTAAAGGCAATGGGAGAGACCTTTTGAAGTC 718
QY 612 AAAGTACTGAAGAGGAAGATTTCATGTTGATCAAGTTACTGTCACAAAGTTTCCAATG 671
Db 719 AAGGACACCGAAGAGAGACTTCCACGTGGACAGGTGACCAACCGTGAAGTGGCTATG 778
QY 672 ATGAAAGACTGGGTATGTTCAATATTAACATATGCAAAATTAAGTCTTTGGGTCTTA 731
Db 779 ATGAAGCGTTTAGGATGTTTAACTCCAGCACTGTGAAGACTGTCGAGCTGGGTGCTG 838
QY 732 TTAATGAAGTATTAGGTAAAGCTACTGCTATTTTATTTTATTTTACAGAGAGAGGTTAAGCTT 791
Db 839 CTGATGAATACTGGGCANTGCCAGCCATCTCTTCTCCTCCCTGATGAGGGGAAACTA 898
QY 792 CAACATTTAGAGATGAGTTGACTCATGACATTTACTTAATTTTATAGAGAACGAGAT 851
Db 899 CAGCAGCTGGAATGAAGTCAACCCACGATATCATCAAGTTCTCTGGAAATGAAGAC 958
QY 852 CGTCGTAGCGCTCTGCACTGCGCAAGTTAAGTATACCGGTACTTACGACTTAAAA 911
Db 959 AGAAGGCTGCGCAGCTTACATTTACCAAACTGTCCATTTACTGGAACCTATGATCTGAAG 1018
QY 912 TCTGTTTTAGGCGAGTTAGGTATTTACCAAGTTTCTTCTACGGTGGCGATTTAGTGGT 971
Db 1019 ACGGCTCGGGTCACTGGGCATCACTAAGGTCTTCAGCAATGGGGCTGACCTCTCCGGG 1078
QY 972 GTTACTGAAGAGCTCCATTTAAATTTGAGTAAAGCTGTTCAAAAGCGCTCTTAAGTATT 1031
Db 1079 GTCACAGAGGAGCCCTGAGCTCTCCAAGGCGGTGCATAGGCTGTGCTGACCAATC 1138
QY 1032 GATGAAAGGGTACCGAGGCGCGCGCTATGTTCTCTGGAAGCTATTCCAATGAGCAT 1091
Db 1139 GACGAGAAAGGACTGAAGCTCTGGGCGCATGTTTCTAGAGGCCATACCCATGCTATC 1198
QY 1092 CCACGAGAGTTAAATTTAATAAACCAATTCGTTTCTCTGATGATGATGATGATGATGAT 1151
Db 1199 CCCCCGAGGTCAAGTTCACAAACCCCTTGTCTCTTAATGATGATGATGATGATGATGAT 1258
QY 1152 AGCCCATTTGTTATGGGTAAAGTTGTCACACCACTCAGAA 1192
Db 1259 TCTCCCTGTTTCATGGGAAAGTGTGATCCCAACCCRAAA 1299

RESULT 5

LOCUS I02706 1312 bp ss-DNA linear PAT 21-MAY-1993
DEFINITION Sequence 1 from Patent US 4599311.
ACCESSION I02706
VERSION I02706.1 GI:268359
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 1312)
Kawasaki, G.H.
TITLE Glycolytic promoters for regulated protein expression: protease inhibitor
JOURNAL Patent: US 4599311-A 1 08-JUL-1986;
1547 - 16th Ave. East, Seattle, WA
FEATURES
Location/Qualifiers
Source 1..1312
BASE COUNT 339 a 368 c 324 g 281 t
ORIGIN
Query Match 28.3%; Score 432.2; DB 6; Length 1312;
Best Local Similarity 60.4%; Pred. No. 5.5e-82;
Matches 713; Conservative 0; Mismatches 468; Indels 0; Gaps 0;

QY 12 GAAGACCTCAAGCGGCGCTCAAAARACCCACAGCTCATCAGACCAAGACCAT 71
Db 100 GAGGATCCCCAGGAGATGCTGCCCAAGAGACAGATACCCACCATGATAGGATCAC 159
QY 72 CCGACTTTTAAATAAATTTACTCCAAATTTAGCCGAATTTGCTTTTCTTTTGTATAGACAA 131
Db 160 CCAACCTTCAACAGATACACCCCAACTTGGCTGAGTTGCGCTTCAGCCTATACCGCCAG 219
QY 132 TTAGCTCATCAAGATAATTTCTACTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTT 191
Db 220 CTGGCACACCAAGTCCAAACAGACCAATATCTTCTCTCCAGTGAGCATCGCTACAGCC 279
QY 192 TTCGCCATGTTAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTTTAGAAGGTTTA 251
Db 280 TTTGCAATGCTCTCCCTGGGACCAAGCTGACACTCAGCATGAAATCTCTGGAGGGCGT 339
QY 252 AACTTTTAAATTTGACCGAAATCCCAAGAGCCCAATTTCAAGGGTTTTCAAGAGTTTCTTG 311
Db 340 AATTTCAACCTCACGAGATTCGAGGCTCAGATCCATGAGGCTTCCAGGAACTCCTC 399
QY 312 AGAATTTGAATCAACCTGATTTCTCAATTTGCAATTAATTAATTAATTAATTAATTAATTA 371
Db 400 CGTACCCTCAACCAAGCCAGACAGCTCCAGCTGACCTGACCCAGGCAATGGCTTCTCCTC 459
QY 372 TCTGAAGTTTAAATTTGGTTGACAAATTTCTAGAACGCTCAAGAACTATATCATAGT 431
Db 460 AGCGAGGCGCTGAAAGCTAGTGGATAAGTTTGGAGGATGTTTAAAGTTTGTACCACTCA 519
QY 432 GAGGCTTTTACCGTTAATTTTGGTGATCTAGGAGGCTAAAGAGCTAAAGAGCAATTAATGATT 491
Db 520 GAAGCTTCTACTGTCACATTCGGGACACCGAGAGGCGCAAGAAACAGATCAAGATTAC 579
QY 492 GTTGAGAAAGGACCCAGGCTTAAGATCGTTGACCTAGTTAAAGAAATTTAGATCGTGATACC 551
Db 580 GTGAGAAAGGCTACTCAAGGGAATTTGGATTTTGTCAAGGAGCTTGACAGAGACACA 639
QY 552 GTCTTCCGACTAGTTAATTTTTCANAGGTAAGTTGGGACGCTCTCTTCGAGGTT 611
Db 640 GTTTTCTCTGGTGAATTTACATCTCTTTTAAAGGCAATGGGAGAGACCTTTGAAGT 699
QY 612 AAAGATCTGAAGAGGAGATTTTTCATGTTGATCAAGTTACTACTGTCAAGTTCCAATG 671
Db 700 AAGCACACCGAGGAGAGGACTTCCAGCTGACGAGCTGACCCAGTGAAGTTCCTATG 759
QY 672 ATGAAAGACTGGGTATGTTCAATATTTCAACATTTGCAAAATTAAGTTCTTGGGTCTTA 731
Db 760 ATGAAGCGTTTAGGCATGTTTAAACATCCAGCATTTGAAGAGCTGTCCAGCTGGGTGCTG 819
QY 732 TTAATGAAGTATTTAGGTAACGCTACTGCTATTTTATTTTACCAGAGGATTAAGCTT 791
Db 820 CTGATGAATACCTGGGCAATGACCGCCATCTCTTCTCTGCTGATGAGGGGAACATA 879
QY 792 CAACATTTAGAGAAATGAGTTGACTCATGACATTTACTAAATTTTATAGAGAACGAGGAT 851
Db 880 CAGCACCTGGAATGAATCAACCCAGGATATCATCAACCAAGTTCTCTGGAATGAAGAC 939
QY 852 CGTCGTAGCGCTTCTCTGCACTGCGCAAGTTAAGTATCAACCGGTACTTACGACTTAAAA 911
Db 940 AGAAGGCTCTCCAGCTTACATTTACCCAACTGTCCAACTACTGGAACCTATGATCTGAAG 999
QY 912 TCTGTTTTAGGCGAGTTAGGTATTTACCAAGTTTCTTAACGCTGCCGATTTGAGTGGT 971
Db 1000 AGGCTCTAGTCACTGGGCATCACTAAGTCTTCAAGCAATGGGCTGACCTCTCCGGG 1059
QY 972 GTTACTGAAGAGCTCCATTTAAATTTGAGTAAAGCTGTTTCAAAAGCCGCTCTTAAGTATT 1031
Db 1060 GTCACAGAGGAGCACCCCTGAAAGCTCTCCAAGGCGGTGATTAAGGCTGTGCTGACCATC 1119
QY 1032 GATGAAAGGGTACCGAGGCGCGGCGGTATGTTCTTGGAGACTATTTCAATGAGCAT 1091
Db 1120 GACGAGAAAGGAGCTGAAGCTGTGGGGCAATGTTTTTAGAGGCCATACCCATGTCTATC 1179

QY 1092 CCACGAGAAGTAAATTAATAAACCAATTCGTTTTCTGTGATCGAGCAGACACTAA 1151
 Db 1180 CGCCCGAGGTCAAGTTCACAAACCCCTTGTCTCTTAATGATGAACAATAACCAAG 1239
 QY 1152 AGCCCATGTTTATGGTGAAGTTGTCAACCAACTCAGAA 1192
 Db 1240 TCTCCCTCTTCAATGGGAAAGTGTGAATCCCAACCAAAA 1280

RESULT 6
 BC011991 1584 bp mRNA linear PRI 02-AUG-2001
 LOCUS Homo sapiens, similar to serine (or cysteine) proteinase inhibitor,
 DEFINITION clade A (alpha-1 antiproteinase, antitrypsin), member 1, clone
 MGC:9222 IMAGE:3859644, mRNA, complete cds.
 ACCESSION BC011991
 VERSION BC011991.1 GI:15080498
 KEYWORDS MGC.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1584)
 Strausberg, R.
 Direct Submission
 Submitted (30-JUL-2001) National Institutes of Health, Mammalian
 Genome Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/DTF
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: villalon@bcm.tmc.edu
 Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
 A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
 Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 21 Row: d Column: 6
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 6855600.
 Location/Qualifiers
 1..1584
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MGC:9222 IMAGE:3859644"
 /tissue_type="Ovary, adenocarcinoma"
 /clone_lib="NIH_MGC_66"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 233..1489
 /codon_start=1
 /product="similar to serine (or cysteine) proteinase
 inhibitor, clade A (alpha-1 antiproteinase, antitrypsin),
 member 1"
 /protein_id="AAH11991.1"
 /db_xref="GI:15080498"
 /translation="MPSSVSWGILLLAGLCLVPVSLAEDPQGDAAQKDTFSHRDQDH
 PTFNKITNLAEFAFSLYROLAHSNSTNIFPSVSTATAFAMLSLGTADKDFLEDKK
 GLNFNLTEPERQHEGQFELLRLTNQDLSQLQTLTGNGLEFLSLKLVDFLEDKK
 LYHSAFTVGTDEEAKQINDYVEKGTQKQIDLVYKELDRDVFALVNIYFPGKW
 ERPFVKQTEEDFVDQVTVKPMKRLGMFIQHCCKLSSWMLVNLGNATALF
 FLPDGKQLHLENEETHDITKFLNEDRRSASLHLPKLSITGYDLKSLVGLGILTK
 VPSNAGDLSGVTEAPLKLKSAVHKAVLTIDEKGTAAAGAMFLEAIPMSIPPEVKFNK
 PFVFLMIDQNTKSPFLFMGKVNPTQK"

FEATURES
 source

CDS

BASE COUNT 410 a 459 c 379 g 336 t
 ORIGIN
 Query Match 28.2%; Score 430.6; DB 9; Length 1584;
 Best Local Similarity 60.3%; Pred. No. 1.2e-81;
 Matches 712; Conservative 0; Mismatches 469; Indels 0; Gaps 0;
 QY 12 GAAGACCCCTCAAGGCGACGCCCTCAAAAACCGACACAGTCATCAGACCAAGACCAT 71
 Db 305 GAGGATCCCGGAGATGCTGCCAGAGAGAGATACATCCACCATGATCAGATCAC 364
 QY 72 CGGACCTTTTAAATAAATAATCTCCAAATTTAGCGGAATTTGCTTTTCTTTGTATAGCAA 131
 Db 365 CCACCTTCAACAGATCACCCCAACCTTGGCTGAGTTCGCCTTCAGCCTATACCGCCAG 424
 QY 132 TTAGCTCATCAAGAGTAATTTCTACTAACAATTTTATTTAGTCTCTGTTTCTATTGCCACTGCT 191
 Db 425 CTGGCACACCAAGTCCCAACAGCACCAATATCTTCTTCCGCCAGTGAGCATCGCTACAGCC 484
 QY 192 TTCGCCATGTTGAGTTTAGGTACTAAAGCCGATACCCATGACGAGATTTTAGAAGGTTTA 251
 Db 485 TTTGCAATGCTCTCCCTGGGGACCAAGGCTGACATCCATGAAGGCTTCAGGAACCTCCCTC 544
 QY 252 AACTTTAAATTTGACCGAAATCCCGAAGCCCAAAATTCACGAGGTTTTCAGAGTTGTTG 311
 Db 545 AATTTCAACCTCAGGAGATTCGGGAGGCTCAGATCCATGAAGGCTTCAGGAACCTCCCTC 604
 QY 312 AGAATTTGAACTCAACTGATTTCTCAATTCGAATTAACCTACTGTTAAACGTTTATTTTGG 371
 Db 605 CGTACCTCTCAACGAGCCAGACAGCCAGCTCCAGCTGACCACCGGCAATGGCTGTTCCTC 564
 QY 372 TCTGAAGGTTTAAATTTGGTTGACAAATTCCTAGAGAGAGTCAAGAAACTATATATAGT 431
 Db 665 AGCGAGGCTCAGCTAGTGGATAGTTTTCGGAGGATTTTAAAGTTGTACCACTCA 724
 QY 432 GAGGCTTTTACCGTTAAATTTGGTGATCTGAGAGAGCTTAAAGCAAAATTAATGATTAT 491
 Db 725 GAAGCTTCACTGTCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTAC 784
 QY 492 GTTGAGAAAGGCGACCCAGGTAAGTCTGTTGACCTAGTTTAAAGAAATTAGATCTGTATAC 551
 Db 785 GTGAGAGAGGCTACTCAAGGGAAATTTGTGATTTGGTCAAGGAGCTTGACAGAGACACA 844
 QY 552 GTCTTCACCTAGTTAACTATATATTTTCAAGGGTAAGTGGGAACGTCTCTTCGAGGTT 611
 Db 845 GTTTTTCCTCTGTGATTTACATCTTCTTAAAGGCAATGGGAGAGACCTTTTGAAGTC 904
 QY 612 AAAGATACCTGAAGAGGAAGATTTTCATGTTGATCAAGTTACTACTGTCAAGTTCCAATG 671
 Db 905 AAGGACACCGAGGAAGAGGACTTCCACGTGGACCGAGGTGACCCCGGTGAAGGTGCCTATG 964
 QY 672 ATGAAAAGACTGGGTATGTTCAATATTTCAATTTGCAAAATAATTAAGTTCTTGGGTCTTA 731
 Db 965 ATGAAGCGTTTTCAGCATGTTTAACTCCAGCTGTGAAGAGCTGTCCACCTGGGTGCTG 1024
 QY 732 TTAATGAAGTATTTAGTAAAGCTGCTATTTTTCCTATTTTACACAGACGAGGTAAAGCTT 791
 Db 1025 CTGATGAATACTCTGGCAATGGCCACCCCATCTCTTCTCTGCTGATGAGGGGAAACTA 1084
 QY 792 CAACATTTAGAGAATAGTTGACTCATGACATTAATTAATTTTATAGAACAGGAGAT 851
 Db 1085 CAGCACCTGGAAAATGAATCACTCACCCAGGATATCATCAAGTTCTCTGGAAAATGAAGAC 1144
 QY 852 CGTGTGACGCTTCTCTGACCTCGCAAGTTAAGTATCACCGGTACTTACGACTTAAAA 911
 Db 1145 AGAAGGCTGCGCAGCTTACATTTACCCAACTGTCCATTTACTGGAACCTATGATCTGAAG 1204
 QY 912 TCTGTCTTAGGCCAGTTAGGTATTACCAAGTTTCTTAAACGGTGGCGATTTGAGTGGT 971
 Db 1205 AGCGTCTGGGTCACTGGGCATCACTAAGGTCTTCAGCAATGGGGCTGACCTCTCCGGG 1264
 QY 972 GTTACTGAAGAGCTCCATTTAAATTTGAGTAAAGCTGTTCAAAAGCGGCTTAACTATT 1031

Best Local Similarity 60.2%; Pred. No. 2.6e-81; Matches 711; Conservative 0; Mismatches 470; Indels 0; Gaps 0;			
QY	12	GAAGACCTCAAGCGGACGCGCTCAAAAACCGACACAGTCAATCACGACCAAGACCAT	71
Db	92	GAGGATCCCGAGGAGATGCTGCCAGACAGATACATCCACCATGATCAGGATCAC	151
QY	72	CCGACTTTAATAAAATTAATCTCAAAATTTAGCCGAATTTGCTTTTCTTTGTATAGACAA	131
Db	152	CCACCTTCACAGATCACCCCAACCTGGCTGAGTTCGCCCTTACGCTATACCCGAC	211
QY	132	TTAGCTCATCAAGTAATCTACTAAATTTTTTTTGTAGTCTGTTCTATTGCCACTGCT	191
Db	212	CTGGCACACAGTCCACAGCAGCAATATCTCTCTCCCGAGTGAGCATCGTACAGCC	271
QY	192	TTCCGCTATGAGTTAGTTAGTAAAGCCGATACCCATCAGCAGATTTTAGAAGGTTTA	251
Db	272	TTTGAATGCTCTCCCTGGGGACCAAGCTCAGACTCAGATGAATCTCGTAGGGCGTG	331
QY	252	AACCTTTAATTTGACCGAAATCCAGAGCCCAATTCACGAGGGTTTTTCAAGAGTTGTTG	311
Db	332	AATTTCAACCTCAGGAGATTCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACCTCTC	391
QY	312	AGAACTTTGAATCAACCTGATCTCAATTTGCAATTAATCTACTGTTAGCGTTTATTTTG	371
Db	392	CGTACCTCTCAACACGACAGCAGCTCCAGCTGACCTACCGCAATGGCCTGTTCCCTC	451
QY	372	TCGAGGTTTAAATGTTGACAAATTCCTAGAACGCTCAAGAACTATATCATAGT	431
Db	452	AGCGAGGCGCTGAAGCTAGTGGATGAATTTTGGAGGATGTTAAAGATGTTGACCACTCA	511
QY	432	GAGGCTTTTACCGTTAATTTTGGTATCTAGGAGCTTAAAGCAAAATTAATGATAT	491
Db	512	GNAGCTTCACTGTCACTTCGGGACACCGAGGCAAGCAAGCAAGATCAACGATTA	571
QY	492	GTGGAAGAGGCCACCGGTAAGATGTTGACCTAGTTTAAAGATTTAGATCGTGATACC	551
Db	572	GTGGAGAAGGGTACTCAAGGGAATTTGGATTTGGTCAAGGAGCTTGACAGAGACACA	631
QY	552	GTCCTCGCAGTATTAATATATTTTCAAGGTAAGTGGGACGTCCTTTTCGAGGTT	611
Db	632	GTTTTGTCTGGTGAATTTACATCTCTTTTAAAGCAATGGGAGACCTTTGAAGTC	691
QY	612	AAAGATCTGAAGAGAGAGATTTTCATGTTGATCAAGTTACTACTGTCAAAAGTTCCAA	671
Db	692	AGGACACCGGAGAGAGACTTCCACGTGGACGAGTGACACCGTGAAGGTGCCATG	751
QY	672	ATGAAAGACGGGTATGTTCAATATCAATTTGCAAAATTAAGTTCTTGCGTCTTA	731
Db	752	ATGAAGCGTTAGGCATGTTTAAATCCAGCAGCTGTAAGAGAGCTGTCCAGCTGGTGCTG	811
QY	732	TTAATGAAGTATTTAGTAAAGCTACTGCTATTTTTTTTACCAGAGCAAGGTAAGCTT	791
Db	812	CTGATGAATACCTGGGCAATGCCACCGCATCTCTTCCTGCTGATGAGGGGAACATA	871
QY	792	CAACATTTAGAGAAATGAGTTGACTCATGACATTAATTAATTTTATAGAAACGAGAT	851
Db	872	GAGCCTTGGAAATGAAGTCAACCAACGATATCATCACCAGTTCTCTGGAATAAGAC	931
QY	852	CGTGTAGCGCTTCTCTGACCTGCCAAAGTTAAGTATACCGGTACTTACGACTTAAA	911
Db	932	AGAAGCTGTGCCAGCTTACATTTTACCCAACTGTCCATTTCTGGAACCTATGATCTGA	991
QY	912	TCGTGTTTATAGGCAATAGTATTAACAAATTTTTTTTACGGTGGCGATTTGAGTGGT	971
Db	992	AGCGCTCTGGGTCACTGGGCATCACTAGGCTCTTACAGATGGGGCTGACCTCTCCGG	1051
QY	972	GTTACTGAAGAGAGCTCCATTAATAATTTAGTAAAGCTGTTTCAAAAGCGCTTTAACTAT	1031
Db	1052	GTCACAGAGGAGCCCTGAGCTCTCAAGCGGTCATCAAGGCTGTGCTGACCATC	1111
QY	1032	GATGAAAGGGTACGAGGCCCGCGGCTATGTTCTTGGAGCTATTTCCATGAGCAT	1091

Db	1112	GACGAGAAAGGACTGAAGCTGCTGGGCCATGTTTTTAGAGCCATACCATGCTCTATC	1171
QY	1092	CCACGAGAGTTAAATTTAATAAACCATTCGTTTTCATGATCGAGCAGCAACTAAA	1151
Db	1172	CCCCCGAGGTCAGTTCAACAAACCTTTGCTCTTTAATGATTGAACAAATACCA	1231
QY	1152	AGCCCATTTTATGGTAAAGTTGTCACACCACTCAGAA	1192
Db	1232	TCCTCCCTCTTCATGGGAAAGTGGTGAATCCACCCAAA	1272

RESULT 9
HUMALATM 1352 bp mRNA linear PRI 01-NOV-1994
LOCUS Human alpha-1-antitrypsin mRNA, complete cds.
DEFINITION K01396
ACCESSION GI:177828
VERSION alpha-1-antitrypsin; antiprotease; antitrypsin.
KEYWORDS Human liver, cDNA to mRNA, clones PULB1523 [1] and pTG603 [3].
SOURCE Homo sapiens
ORGANISM 1 (bases 1 to 1231)
REFERENCE 1 (bases 1 to 1231)
AUTHORS Bollen, A., Herzog, A., Cravador, A., Herion, P., Chuchana, P., Vander Straten, A., Loriau, R., Jacobs, P. and van Elsen, A.
TITLE Cloning and expression in Escherichia coli of full-length complementary DNA coding for human alpha 1-antitrypsin
JOURNAL DNA 2 (4), 255-264 (1983)
MEDLINE 84107980
PUBMED 6319097
REFERENCE 2 (bases 1 to 1352)
AUTHORS Colau, B., Chuchana, P. and Bollen, A.
TITLE Revised sequence of full-length complementary DNA coding for human alpha 1-antitrypsin
JOURNAL DNA 3 (4), 327-330 (1984)
MEDLINE 85026667
PUBMED 633329
REFERENCE 3 (bases 95 to 286)
AUTHORS Courtney, M., Buchwalder, A., Tessier, L.H., Jaye, M., Benavente, A., Balland, A., Kohli, V., Lathé, R., Toistoshev, P. and Lecocq, J.P.
TITLE High-level production of biologically active human alpha 1-antitrypsin in Escherichia coli
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 81 (3), 669-673 (1984)
MEDLINE 84144765
PUBMED 6322161
COMMENT Alpha-1-antitrypsin is synthesized in the liver and is a major constituent of plasma. It functions as an inhibitor of elastase, which degrades lung tissues. Hereditary deficiencies are fairly frequent, causing early lung degeneration and sometimes severe liver disorders. Defectives are particularly sensitive to pollution induced inflammation. Injection of anti-trypsin relieves the condition [1].

FEATURES	Location/Qualifiers
source	1. .1352 /organism="Homo sapiens" /db_xref="taxon:9606" /map="14q32.1"
gene	1. .1352 /gene="PI"
mRNA	<1. .1352 /gene="PI" /product="al-antitrypsin mRNA"
CDS	20. .1276 /gene="PI" /note="alpha-1-antitrypsin" /codon_start=1 /protein_id="AAB59375.1" /db_xref="GI:177828" /db_xref="GBD:G00-120-289" /translation="MPSSVSWGILLIAGLCCILPVSVLAEDPQDAAQKTDTSHHDDH PTNKTPNLAFAFSLYQALQHSNSTIFFSPVSIATAFAMLSIGTKADHDELE GLNFNTEIPEAQIHGFEQLLNTQDLSQLTGNGLFLSEGLKLVDFKLEFVKK LYHSEFTVNFQDTEAKQINDYVEKQGTQKIVDLVKELDRDRTFVLYNFFKQK

BRPPEVKDEERDHFDOVTTVKVPMKRLGFMFNIOHCKKLSSHVLLMKYLGNATAIF
FLPDGKQLHLENELTHDIITKFLNEDRRSASHLPLKLSITGYDLKSVLUGUGITK
VFSNGADUSGVTEERAPLKLRAVHKAVLTIDKEKTEAAGAMFLEAIPMSIPPEVKFNK
PFVFLMIQNTKSPLEFMKGKVVNPTQK*
20. .91

sig_peptide

/gene="pi"
/note="alpha-1-antitrypsin signal peptide"
92. 1273
/gene="pi"
/product="alpha-1-antitrypsin mature peptide"
241
/gene="pi"
/citation=[1]
286
/gene="pi"
/citation=[3]
/replaces="*"
435. 439
/gene="pi"
/citation=[1]
748
/gene="pi"
/citation=[1]
784
/gene="pi"
/citation=[1]
787
/gene="pi"
/citation=[1]
799. 802
/gene="pi"
/citation=[1]
837
/gene="pi"
/citation=[1]
1114
/gene="pi"
/citation=[1]

old_sequence

conflict

old_sequence

old_sequence

old_sequence

old_sequence

old_sequence

old_sequence

old_sequence

BASE COUNT 349 a 386 c 325 g 292 t
ORIGIN 90 bp upstream of Ddel site.

Query Match 28.1%; Score 429; DB 9; Length 1352;

Best Local Similarity 60.2%; Pred. No. 2.6e-81;

Matches 711; Conservative 0; Mismatches 470; Indels 0; Gaps 0;

QY 12 GAAGACCTCAAGGCGACCGCTCAAAAACCGACACAGTCATCAGACCAAGACCAT 71

DB 92 GAGGATCCCGGGAGATGCTGCCAGAGACAGATACATCCACCAGATCAGATCAC 151

QY 72 CCGACTTTTAAATAAATACTCCAATTTAGCCGAATTTGCTTTTCTTTGATAGACAA 131

DB 152 CCACCTTCAACAGATCACCCCAACCTGGCTGAGTTGCCCTCAGGCTATACGCCAG 211

QY 132 TTAGCTCATCAAGTAATTTCTACTAACATTTTTTTTAGTCTGTGTTCTATTGCCACTGCT 191

DB 212 CTGGCACACCTCCCAACAGCAATATCTTCTCTCCAGTGAGCATGCTACAGCC 271

QY 192 TTCGCCATGTTAGTTAGTACTAAGCCGATCCCATGACGAGATTTTGAAGGTTTA 251

DB 272 TTGCAATGCTCTCCCTGGGACCAAGGCTGACACTCAGATGAATCTGAGGCGCTG 331

QY 252 AACTTTAATTTGACCGAATCCAGAGCCCAATTTACAGAGGTTTTCAGAGTTGTTG 311

DB 332 AATTTCAACCTCAGGAGATTCGGAGGCTCAGATCCCATGAAGGTTCCAGGAATCCTC 391

QY 312 AGAAGCTTTGAATCAACCTGATTTCTCAATTTGAATTTTAACTACTGGTAAGGTTTATTG 371

DB 392 CGTACCTCAACAGCCAGACAGCTCCAGCTGACCCAGGCAATGGCTGTCTC 451

QY 372 TCTGAAGTTTAAATTTGTTGACAAATTCCTAGAACGCTCAAGAACTATATAGT 431

DB 452 AGCGAGGCGCTGAAGCTAGTGAATTTTGGAGGATGTTAAAAAGTTGTACCACCTCA 511

QY 432 GAGGCTTTTACCGTAAATTTTGGTGATCTGAGGAAGCTAAAAGCAAAATTAATGATTAT 491

DB 512 GAAGCCTTCACTGTCAACTTTCGGGGACACCGAAGAGGCCAAGAACAGATCAAGGATTAC 571

QY 492 GTTGAGAAAGCCACCCAGGTAAGATCGTTGACCTAGTTAAAGAAATTTAGATCGTGATACC 551

DB 572 GTGAGAAAGGTACTCAAGGGAATAATTTGGTCAAGGAGCTTGACAGAGACACA 631

QY 552 GTCCTCCACTAGTAACTATATTTTTTCAAGGGTAAGTGGGAACGCTCTTTTCAGAGTT 611

DB 632 GTTTTTGCTGTGGTAATTTACATCTTCTTTAAAGGCAATTTGGGAGACACCTTTTGAAGTC 691

QY 612 AAGATCTGAAGAGGAAGATTTTCATGTGTGATCAAGTTACTACTGTCAAGTTCCAATG 671

DB 692 AAGACACCCGAGGAAGAGGACTTCCAGTGGACAGGTGACCCGTAAGAGGTGCGCTATG 751

QY 672 ATGAAAGACTGGGTATGTTCAATATTCAACATTTCAAAAAAATTTAAAGTTCTTGGGTCTTA 731

DB 752 ATGAGCGTTTAGGCATGTTTAACATCCAGCACTGTAAAGAGCTGTCCAGCTGGGTCTG 811

QY 732 TTAATGAAGTAATTTAGGTAAGCTACTGTGCTATTTTTTTTTTACCAGACGAGGTAAGCTT 791

DB 812 CTGATGAATACCTGGGCAATGCCACGCCATCTTCTTCTGCTGATGAGGGGAAACTA 871

QY 792 CAACATTTAGAGTAAGTTGACTCATGACATTTATTAAATTTTATAGAGACGAGGAT 851

DB 872 CAGCACCTGGAAATGAACCTCACCCAGATATCATCCCAAGTTCTTGGAAATGAAGAC 931

QY 852 CGTGTAGCGCTTCTCTGCACCTGCCAAAGTTAAGTATCACCGGTACTTACGACTTAAAA 911

DB 932 AGAAGGTCTGCCAGCTTACATTTACCCCAACTGTCCATTACTGGAACCTATGATCTGAAG 991

QY 912 TCTGTTTTAGGCCAGTTAGGTATTACCAAGTTTTTTTCTAACGGTCCGATTTTGAAGTGT 971

DB 992 AGCGTCTGGTCAACTGGGCATCACTAAGGTCTTTCAGCAATGGGGTGACCTCTCGGG 1051

QY 972 GTTACTGAAGAAGCTCCATTAAAAATGAGTAAGCTCTTCACAAAGCCGCTTAACCTATT 1031

DB 1052 GTCAGAGGAGGCCCCCTGAGCTCTCCAGGCCGTGATAGGCTGTGCTGACCATC 1111

QY 1032 GATGAAAGGTTACCGAGCGCCGCGCTATGTTCTTGGAAAGCTATTCGAATGAGCAT 1091

DB 1112 GACGAGAAAGGACTGGAAGCTGTGGGCCATGTTTTTAGAGGCCATACCATGCTCTATC 1171

QY 1092 CCACCAAGATTTAAATTTAATAAACCAATTCGTTTTCTGTATGATCGAGGAGACACTAAA 1151

DB 1172 CCCCCGAGGTCAAGTTCAACAAACCTTTTGTCTTCTTAATGATTGAACAAAATACCAAG 1231

QY 1152 AGCCATTGTTTATGGGTAAGTTGTCAACCCCAACTCAGAA 1192

DB 1232 TCTCCCTCTTCATGGGAAAAGTGGTGAATCCCAACCCAAAA 1272

RESULT 10

AX335338

LOCUS

DEFINITION

AX335338

ACCESSION

AX335338.1

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

AX335338 Sequence 5847 from Patent WO0194629. 1371 bp DNA linear PAT 09-JAN-2002

AX335338

AX335338

AX335338.1 GI:18126057

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1

Young, P. E., Augustus, M., Carter, K. C., Ebner, R. R., Endress, G.,

Horrigan, S., Soppet, D. R. and Weaver, Z.

Cancer gene determination and therapeutic screening using signature

gene sets

Patent: WO 0194629-A 5847 13-DEC-2001;

Avalon Pharmaceuticals (US)

Location/Qualifiers

1. .1371

```
Query Match 28.1%; Score 429; DB 6; Length 1371;
Best Local Similarity 60.2%; Pred. No. 2.6e-81;
Matches 711; Conservative 0; Mismatches 470; Indels 0; Gaps 0;

BASE COUNT 350 a 388 c 340 g 293 t
ORIGIN

QY 12 GAAGACCTCAAGCGGAGCGCTCAAAAACCGACACAGTATCATACGACCAAGACCAT 71
DB 111 GAGGATCCCCAGGAGATGCTGCCAGAGACAGATATATCCACCAATGATCAGGATCAC 170
QY 72 CCGACTTTTAATAAATTAATCTCCAAATTTAGCCGAATTTGCTTTTCTTTTGTATAGACAA 131
DB 171 CCAACCTTCAACAGATCAACCCCACTGCTGAGTTCGCTTTCAGCCTATACCGCCAG 230
QY 132 TTAGCTCATCAAGTAATCTTACTAATATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 191
DB 231 CTGGCACACCAAGTCCACAGACCAATATCTTCTTCCCGAGTGAGCATCGCTACAGCC 290
QY 192 TTCGCATGTTGAGTTAGGTAAGGCGATACCCATGACGAGATTTTGAAGGTTTA 251
DB 291 TTGCAATGCTCTCCCTGGGACCAAGCTGACACTCAGATGAATCCTTGGAGGCGCTG 350
QY 252 AACTTTAATTTGACGAAATCCCAAGACCCCAATTCACGAGGTTTTCAGAGATTTGTTG 311
DB 351 AATTTCAACCTCAGGAGATTCGGAGGCTCAGATCCATCAAGGCTTCCAGGAACCTCCTC 410
QY 312 AGAATTTGAATCAACCTGATCTCAATGCAATTAATCTACTGTAACGGTTTATTTTG 371
DB 411 CGTACCTTCAACCAAGCCAGACGAGCTCCAGCTGACCCGCAATGGCTTCTCCTC 470
QY 372 TCTGAAGTTTAAATTTGGTTGACAAATTCCTAGAAGCGTCAAGAACTATATCATAGT 431
DB 471 AGCGAGGCTTGAAGCTAGTGTGATAGTTTGTGAGATGTTAAAGATGTTGACCACTCA 530
QY 432 GAGGCTTTTACCGTTAATTTTGTGATCTAGTGAAGCTCAAAAGCAAAATTAATGATTAT 491
DB 531 GAAGCTTCTACTGTCACTTCCGGGACACCGAGGCGCAAGAGGCAAGAAACAGATCAAGATAC 590
QY 492 GTTGAGAAGCCACCCAGGTTAAGATCGTTGACCTAGTTAAAGATTAAGATCGTGATACC 551
DB 591 GTGGAGAAGGTTACTCAAGGAAATTTGTGATTTGTGCAAGGAGCTTGACAGAGACACA 650
QY 552 GTCTTCGCACGTAGTAACTATATTTTTCAGGGTAAAGTGGGACGCTCTTTCGAGGTT 611
DB 651 GTTTTGTCTGTTGAATTAATCTCTTTTAAAGCAAAATGCGAGAGACCTTTTGAAGTC 710
QY 612 AAAGATCTGAAGAGGAGATTTTCATGTTGATCAAGTTACTACTGTCAAAAGTTCCAAATG 671
DB 711 AAGGACACCGAGGAGAGGACTTCCAGTGGACCAAGTGCACCCGTTGAAGGTGCCTATG 770
QY 672 ATGAAAAGACTGGGTATGTTCAATATTCACATTTGCAAAATTAAGTTCTTGGGTCTTA 731
DB 771 ATGAGCGTTTAGCATGTTTAAATCATCAGCACTGTAAAGAGCTGTCCAGCTGGGTGCTG 830
QY 732 TTAATGAAGTATTTAGGTAAGCGTACTGCTATTTTTTTTTTACCAGAGGAGTAAAGCTT 791
DB 831 CTGATGAATACCTGGGCAATGCCACCCCATCTTCTTCTCCCTGATGAGGGGAAACTA 890
QY 792 CAACATTTAGAGAAGTCTGACTCATGACATTTACTAATTTTATAGAAACGAGGAT 851
DB 891 CAGCACCTTGAAGAAATGAATCAACCCACGATATCATCAAGATTCCTGGAAATTAAGAC 950
QY 852 CGTCCGTAGCGTTCTCTGACCTGCCAAAGTTAAGTATCACCGGTACTTACGACTTAAAA 911
DB 951 AGAAGGCTGCCAGCTTACATTTACCAAACTGCTTCTTACTGGAACCTATGATCTGAAG 1010
QY 912 TCTGTTTATAGCCAGTTAGGTATTAACCAAGTTTCTTCTTACCGGTGCGGATTTGAGTGGT 971
DB 1011 AGCGTCTCTGGGTCAACTGGGATCACTAAGGCTTCTAGCAATGGGGTGTGACCTCTCGGG 1070

/organism="Homo sapiens"
/db_xref="taxon:9606"
293 t
340 g
388 c
350 a
```

LYHSAFTVNFEDTBEAKKQINDYVEKGTQKIVDLVKELDRDVTVALVNYIFFKGM
ERPFVKQTEEDFHDVITVKVPMKRLGMENIOHCKKLSSWYLLKVLGNATAIF
FLPDEKGLHLENELTHDITKELENERDRSALHLPKLSITGYDLKSVLQGLGTR
VFSNADLSGVTEERAPLKSAYKHAVLITDEKGTAEAGAMFLAIPMSIPEPVKFNK
PFVFLIMISONTKSPFLFMGVNPNPQK*
39..110
111..1292
/product="unnamed"

sig_peptide
mat_peptide

polyA_site 350 a 388 c 340 g 293 t

1371

BASE COUNT
ORIGIN

Query Match 28.1%; Score 429; DB 9; Length 1371;

Best Local Similarity 60.2%; Pred. No. 2.6e-81;

Matches 711; Conservative 0; Mismatches 470; Indels 0; Gaps 0;

QY 12 GAAGACCCCTAAGGCGACCGCGTCAAAAACCGACACGATCATCACGACCAAGACCAT 71

DB 111 GAGGATCCCGAGGAGATGTCGCCAGAGACAGATACATCCACCATGATCAGGATCAC 170

QY 72 CGGACCTTTAATAAATAATCTCCAATTTAGCGGAATTTGCTTTTCTTTGTATAGACAA 131

DB 171 CCAACCTTCAACAGATACACCCCAACCTGGCTGAGTTCCGCTTCAGCGTATATACGCCAG 230

QY 132 TTAGCTCATCAAGTAATCTACTAATATTTTCTAGTCTGTTTCTATTGCCACTGCT 191

DB 231 CTGGCACACAGTCCACAGCACCATATCTTCTTCTCCCGAGTGAGCATCCGACGCC 290

QY 192 TTGCGCATGTGATTTAGTACTAAGCCGATACCCATGACGAGATTTTGAAGGTTTA 251

DB 291 TTTGCAATGCTCTCCCTGGGGACCAAGGCTGACACTCAGCATGAAATCTCGGAGGCGCTG 350

QY 252 AACTTTAATTTGACCAATCCAGAGCCCAATTCACGAGGGTTTCAAGATTTGTTG 311

DB 351 AATTTCAACCTCACGGAGATCCGGAGGCTCAGATCCATGAGGCTTCAGGAACCTCTC 410

QY 312 AGAATTTTGAATCAACCTGATTTCAATTCGAATTAACCTAGTGGTAACGGTTTATTGTTG 371

DB 411 CGTACCTTCAACAGCCAGACAGCCAGCTCCAGCTGACCAACCGCAATGGCTGTCTCTC 470

QY 372 TCTGAAGGTTTAAATTTGTTGACAAATTTCTTAGAAGAGCTCAAGAAATATATCATAGT 431

DB 471 AGCGAGGCGCTGAAGCTAGTGGATAGTTTGGAGGATGTTAAAGTTGTACCACTCA 530

QY 432 GAGGCTTTTACCGTTAATTTGTTGATGACTAGGAGCTTAAAGCAATTAATGATAT 491

DB 531 GAAGCTTCTACCTGTCACCTTCGCGGACACCGAGAGGCCAAGAAACAGATCAACGATTAC 590

QY 492 GTTGAGAAAGGCCACCCAGGCTAAGATCGTTGACCTAGTTAAAGAAATAGATCGTGATACC 551

DB 591 GTGAGAGGGTACTCAAGGAAATTTGGATTTGGTCAAGGAGCTTGACAGACACACA 650

QY 552 GTCTTCCGCTAGTTAACTATATTTTTCAGGGTAAAGTGGAGCGTCTCTTCAGGTT 611

DB 651 GTTTTGTCTGTGTGAATACATCTCTTTTAAAGCAATGGAGAGACCCCTTTGAAGTC 710

QY 612 AAGATACTGAAGAGAAATTTTATGTTGATCAAGTTACTACTGCTCAAAAGTTCCAAATG 671

DB 711 AAGGACACCGAGGAAGAGACTTCCAGCTGGACCGAGTGACCCAGCTGAAGGTGCTATG 770

QY 672 ATGAAAGACTGGGTATGTTCAATATTAACATTTGCAAAAATTAAGTTCTTTGGGTCTTA 731

DB 771 ATGAGCGTTTATGCAATGTTTACATCCAGCACTGTGAAGAGCTGTCAGCTGGGTGCTG 830

QY 732 TTATGAAGTATTTAGGTAAACGCTACTGCTATTTTTTTTTTACCAGCAGGAAGTAAAGCTT 791

DB 831 CTGATGAATACCTGGGCAATGCCACCGCATCTTCTTCTGCTGATGAGGGAACTA 890

QY 792 CAACATTTAGAAATGAGTTGACTCATGACATTAATTAATTTTATAGAACAGGAT 851

DB 891 CAGCACTTGGAAATGAATCAACCCAGCATATCATCACCAGTTCTCTGGAATAAGAC 950

QY 852 CGTCTGAGCGCTCTCTGCACTGCGCAAGTTAAGTATCACCGGTACTTACGACTTAAA 911

Db 951 AGAAGGCTGCCAGCTTACATTTACCACAACTGTCCATTACTGGAACCTATGATCTCAAG 1010

Qy 912 TCTGTTTTAGCCAGTAGTATTAACAAAGTTTTTTCTTAACGGTGGCGGATTTGAGTGGT 971

Db 1011 AGCGTCTGGTCAACTAGGCGATCACTAAGGCTCTTCAGCAATGGGCGTCCGCTCCGGG 1070

Qy 972 GTTACTGAAGAAGCTCCATTAATAATTTAGTAAAGCTGTTCACAAGCGCTTTAACTATT 1031

Db 1071 GTCACAGAGGAGGCCCTGAACTCTCCNAGGCCGTGCATAGGCTGTGCTGACCATC 1130

Qy 1032 GATGAAAGGGTACCGAGCGCGCGGCTATGTTCTTGGAGCTATTTCCAATGAGCAT 1091

Db 1131 GACGAGAAAGGAGCTGAAGCTGCTGGGCCATGTTTTTAGAGGCCATACCATGCTCTATC 1190

Qy 1092 CCACGAGAAGTTAATTTAATAAACCATTCGTTTTCTTGATGATCGAGCAGAACACTAAA 1151

Db 1191 CCCCCGAGGTCAAGTTTCAACAAACCCCTTTGCTTCTTAATGATTGAACAAATATACCAG 1250

Qy 1152 AGCCATTTGTTATGGGTAAAGTTGTCAACCCCAACTCAGAA 1192

Db 1251 TCTCCCTCTTCATGGGAAAAGTGGTGAATCCCAACCCAAAA 1291

RESULT 12

AK026174

LOCUS

DEFINITION

AK026174

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

misc_feature

BASE COUNT

ORIGIN

1..1399

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="HRC12386"

/cell_type="primary human renal epithelial cells"

/clone_lib="HRC"

/note="cloning vector pME18SFL3"

1..1399

/note="highly similar to AF113676 Homo sapiens clone

FLB2803 PRO0684 mRNA"

378 a 395 c 329 g 297 t

1399 bp mRNA linear PRI 29-SEP-2000

FLJ22521 fis, clone HRC12386, highly similar to

AK026174

AK026174.1 GI:10438937

oligo capping; fis (full insert sequence).

Homo sapiens primary human renal epithelial cells cDNA to mRNA,

clone_lib:HRC clone:HRC12386.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (sites)

Kawabata,A., Hiki,j,i., Kobatake,N., Inagaki,H., Ikema,Y.,

Okamoto,S., Okitani,R., Ota,T., Suzuki,I., Obayashi,M., Nishi,T.,

Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.

NEDO human cDNA sequencing project

Unpublished

2 (bases 1 to 1399)

Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,

Shibahara,T., Tanaka,T. and Nakamura,Y.

Direct Submission

Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,

University of Tokyo, Laboratory of Genome Structure Analysis, Human

Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,

Japan (E-mail:cdnafeims.u-tokyo.ac.jp, Tel:81-3-5449-5286,

Fax:81-3-5449-5416)

NEDO human cDNA sequencing project supported by Ministry of

International Trade and Industry of Japan: cDNA full insert

sequencing: Research Association for Biotechnology: cDNA library

construction, 5'- & 3'-end one pass sequencing: Department of

Virology and Human Genome center, Institute of Medical Science,

University of Tokyo (partly supported by Science and Technology

Agency).

Location/Qualifiers

1..1399

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="HRC12386"

/cell_type="primary human renal epithelial cells"

/clone_lib="HRC"

/note="cloning vector pME18SFL3"

1..1399

/note="highly similar to AF113676 Homo sapiens clone

FLB2803 PRO0684 mRNA"

378 a 395 c 329 g 297 t

Query Match	28.1%;	Score 429;	DB 9;	Length 1399;	
Best Local Similarity	60.2%;	Pred. No. 2.6e-81;			
Matches 711;	Conservative	0;	Mismatches 470;	Indels	0; Gaps
QY	12	GAAGACCCCTCAAGGCGACCGCTCAAAAACCGACACAGCTATCATCAGCAGCAAGACCAT	71		
DB	116	GAGGATCCCAGGAGATGCTGCCAGAGAGACAGATACATCCACCATGATCAGGATCAC	175		
QY	72	CGACATTTAATAAATPACTCCAAATTTAGCCGAATTTGCTTTTCTTTGTATAGACAA	131		
DB	176	CCAACCTTCAACAAGATACCCCCCAACCTGGGTGAGTTCGGCTTCAGCCTATATCTGCCAG	235		
QY	132	TTAGCTCATCAAGTAATTTCTACTAACATTTTTTTTATGCTGTTTCTATTGCCACTGCT	191		
DB	236	CTGGCACACCATGCTCAACAGCACCAATATCTTCTCTCCAGTACAGATCCTACAGCC	295		
QY	192	TTCCGCAATGTGAGTTTGTAGTACTAAAGCCGATACCCATGACGAGATTTTAAAGGTTTA	251		
DB	296	TTTGCATGCTCTCCCTGGGGACCAAGGCTGACATCAGATGAATCCTGGAGGGCGTG	355		
QY	252	AACCTTAATTTGACCGAAATCCAGAGCCCAAAATTCAGAGGGTTTTCAAGAGTTGTTG	311		
DB	356	AATTTCAACTACGGAGATTCGGAGGCTCAGATCCATGAGGCTCCAGGAACCTCCTC	415		
QY	312	AGAACTTTGAATCAACCTGATTTCTCAATTTGCAATTAATCTACTGTPAAGGTTTATTTTG	371		
DB	416	CGTACCTCAACAGCAGACAGCCAGCTCCAGCTGACCGCGCAATGGCGTGTCTCCTC	475		
QY	372	TCCTGAAGTTTAAATTTGCTGACAAATTCCTPAGAGAGCTCAAGAAACTATATCATAGT	431		
DB	476	AGCGAGGGCCCTGAAGTAGTGTGAATGTTTTTGGAGGATGTTAAAGATTTGTACCACTCA	535		
QY	432	GAGGCTTTTACCCTTAATTTTGGTGATCTAGGAGCTTAAAGCAATTAATGATTAT	491		
DB	536	GAAGCCTTCTACTGTCACCTCGGGGACCCGAGAGGCCAAGAAACAGATCAACGATTAC	595		
QY	492	GTGAGAAAGGCCACCGGTAAGATCGTTGACCTAGTTTAAAGAAATTAAGTCGTGATACC	551		
DB	596	GTGGAGAGGGTACTCAAGGGAAATTTGGGATTTGGTCAAGAGCTTGACAGAGACACA	655		
QY	552	GTCTTCGCACCTAGTTAACTATATTTTTTCAAGGGTAAAGTGGGAACGCTCTTTCAGGTT	611		
DB	656	GTTTTGTCTGTGATTAATCTCTTTAAAGCAATGGAGAGACCCCTTTTGAAGTC	715		
QY	612	AAAGATCTAGAGGAAGATTTTCAATGTTGATCAAGTTTACTACTGCTCAAGTTTCCAATG	671		
DB	716	AAGGACACCGAGGAAGGACTTCCACGTGGACAGGTGACCAACCGTGAAGGTGCGCTATG	775		
QY	672	ATGAAGAGCTGGGTATGTTCAATATTCACATTCGCAAAATTAAGTCTTGGGCTCTTA	731		
DB	776	ATGAGGGTTTAGCAATGTTTAACTATCCAGACTGTAAAGAACTGTCCAGCTGGGTGCTG	835		
QY	732	TTAATGAAGTATTAGGTAAAGCTACTGCTATTTTTTTTTTACCAGAGGAAGGTAAAGCTT	791		
DB	836	CTGATGAATACCTGGGCAATGCCACCGCATCTTCTCTCCCTGATGAGGGGAAACTA	895		
QY	792	CACATTTAGAGATGAGTTGACTCATGACATTAATTAATTTTAAAGAACAGGAT	851		
DB	896	CAGCACCTGAAAATGAACATCACCCAGATATCATCAACCAAGTTCTCGTGAATTAAGAC	955		
QY	852	CGTCGTAGCGCTTCTCTGCACTGCGCAAAAGTTAAAGTATCACCGGTACTTACGACTTAAA	911		
DB	956	AGAGGTCTGCCAGCTTACATTTTACCAAACTGTCCATTTACTGGAACCTATGATCTGAAG	1015		
QY	912	TCGTGTTTGGCCAGTTAGGTATTAACAAAGTTTCTTAAAGGTCGCGCATTTGAGTGGT	971		
DB	1016	AGCGTCTGGGTCAACTGGGCACTACTAAGGTCTTTCAGCAATGGGGCTGACCTCTCCGGG	1075		
QY	972	GTTACTGAAGAAGCTCCATTAATTTAGTAAGCTGTTTCAAAAAGCCGCTCTTAATCTATT	1031		
DB	1076	GTCACAGAGGAGGACCCCTGAAAGCTCTCCAGGCGCGTGCATAAAGGCTGTGCTGACCATC	1135		
QY	1032	GATGAAAAGGATCCAGAGGCGCGCGCTATGTTTCTGGAAGCTATTTCATGAGCATTTAGCATT	1091		

QY	192	TTCCGCATGTTAGTTAGTACTAAGCGCATACCATGACGAGATTTTGAAGGTTTA	251
Db	299	TTTGGCAATGCTCTCCCTGGGGACCAAGCTGACACTCAGATGAATCTGGAGGCGTG	358
QY	252	AACCTTTAAATTTGACGAATCCAGAGCCCAAAATTCACGAGGGTTTTCAAGAGTTGTTG	311
Db	359	AATTTCAACCTCACGGAGATTCGGAGGCTCAGATCCATGAAGGCTTCAGGAACCTCC	418
QY	312	AGAACTTTGAATCAACCTGATCTCAATTTGCAATTAACCTACTGTTAACTGTTT	371
Db	419	CGTACCTTCAACGACGACAGCCAGCTCCAGCTGACCCAGGCAATGGCTGTTCC	478
QY	372	TCCTGAAGGTTTAAATTTGGTTGACAAATTCCTAGAAGAGCTCAAGAAACTATATCAT	431
Db	479	AGCGAGGCGCTGAAGCTAGTGATGATTTTGGAGGATGTTAAAGTTGTACCACTCA	538
QY	432	GAGGCTTTTACCGTTAAATTTGGTGATACTGAGGAGCTAAAGCAAAATTAATGATTA	491
Db	539	GAAGCCTTCACTGCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTA	598
QY	492	GTTGAGAAAGGTCACCGAGGTAAGATCGTTGACCTAGTTAAAGAAATTAAGCTGTA	551
Db	599	GTGGAGAAGGTTACTCAAGGGGAAATTTGGATTGGTCAAGGAGCTTGACAGAGACA	658
QY	552	GTCTTCGCACTAGTTAACTATATTTTTTCAAGGGTAAGTGGGAACTGCTTTCAGG	611
Db	659	GTCTTTGCTCTGGTGAATTTACATCTCTTTAAAGCAAAATGGAGAGACCTTTG	718
QY	612	AAAGATCTGAAGAGGAAGATTTTCATCTTGAATCAAGTTACTGCTCAAAATGTTCA	671
Db	719	AAGGACACCGAGGAAGAGGACTTCCACGTGGACCGTGAACCGGTGAAGGTCCTAT	778
QY	672	ATGAAAGACTGGGTATGTTCAATTTCAACATTCACAAATTAAGTTCTTTGGTCTTA	731
Db	779	ATGAAGCGTTTAGGCATGTTTAAACATCCAGCACTGTAAGAAGCTGTCCAGCTGG	838
QY	732	TTAATGAAGTATTTAGGTAAAGCTACTGCTATTTTTTTTTTACCAGAGCAAGGTA	791
QY	192	TTCCGCATGTTAGTTAGTACTAAGCGCATACCATGACGAGATTTTGAAGGTTTA	251
Db	299	TTTGGCAATGCTCTCCCTGGGGACCAAGCTGACACTCAGATGAATCTGGAGGCGTG	358
QY	252	AACCTTTAAATTTGACGAATCCAGAGCCCAAAATTCACGAGGGTTTTCAAGAGTTGTTG	311
Db	359	AATTTCAACCTCACGGAGATTCGGAGGCTCAGATCCATGAAGGCTTCAGGAACCTCC	418
QY	312	AGAACTTTGAATCAACCTGATCTCAATTTGCAATTAACCTACTGTTTAACTGTTT	371
Db	419	CGTACCTTCAACGACGACAGCCAGCTCCAGCTGACCCAGGCAATGGCTGTTCC	478
QY	372	TCCTGAAGGTTTAAATTTGGTTGACAAATTCCTAGAAGAGCTCAAGAAACTATATCAT	431
Db	479	AGCGAGGCGCTGAAGCTAGTGATGATTTTGGAGGATGTTAAAGTTGTACCACTCA	538
QY	432	GAGGCTTTTACCGTTAAATTTGGTGATACTGAGGAGCTAAAGCAAAATTAATGATTA	491
Db	539	GAAGCCTTCACTGCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTA	598
QY	492	GTTGAGAAAGGTCACCGAGGTAAGATCGTTGACCTAGTTAAAGAAATTAAGCTGTA	551
Db	599	GTGGAGAAGGTTACTCAAGGGGAAATTTGGATTGGTCAAGGAGCTTGACAGAGACA	658
QY	552	GTCTTCGCACTAGTTAACTATATTTTTTCAAGGGTAAGTGGGAACTGCTTTCAGG	611
Db	659	GTCTTTGCTCTGGTGAATTTACATCTCTTTAAAGCAAAATGGAGAGACCTTTG	718
QY	612	AAAGATCTGAAGAGGAAGATTTTCATCTTGAATCAAGTTACTGCTCAAAATGTTCA	671
Db	719	AAGGACACCGAGGAAGAGGACTTCCACGTGGACCGTGAACCGGTGAAGGTCCTAT	778
QY	672	ATGAAAGACTGGGTATGTTCAATTTCAACATTCGAAATTAAGTTCTTTGGGTCCTTA	731
Db	779	ATGAAGCGTTTAGGCATGTTTAAACATCCAGCACTGTAAGAAGCTGTCCAGCTGG	838
QY	732	TTAATGAAGTATTTAGGTAAAGCTACTGCTATTTTTTTTTTACCAGAGCAAGGTA	791
Db	839	CTGATGAATACCTGGGCAATGCCACCGCATCTCTCTCGCTGATGAGGGGAACTA	898
QY	792	CAACATTTAGAGAAATGAGTTGACTCATGATATTAATTAATTAATTTTGAAGCAAC	851
Db	899	CAGCAGCTGGAATGAATCACTCAGGATATCATCAACCAAGTTCCCTGGAAATGAAG	958
QY	852	CGTCTGAGGCTTCTCTGACCTGCGCAAGTTAAGTATACCGGTACTTACGACTTAAA	911
Db	959	AGAAGTCTGCCAGCTTACATTTTACCAACTGTCCATTAAGTGAACCTATGATCTGA	1018
QY	912	TCCTGTTTGGCCAGTTAGTATTTACCAAGTTTCTTAAAGGTCGCGATTTGAGTGT	971
Db	1019	AGGCTCTGGTCACTGGGCACTACTAAGGCTCTCAGCAATGGGCTGACCTCTCCGG	1078
QY	972	GTTACTGAAGAGCTCCATTAATTAATTAAGTAAAGCTGTTCAAAAGCCGCTCTTA	1031
Db	1079	GTACAGAGGAGGACCCCTGAGGCTTCCAGGCGCTGCAATGAAGCTGTGCTGACCAT	1138
QY	1032	GATGAAAGGTTACCGAGGCGCGGCTATGTTCTGGAAGCTATTCATGAGCAAT	1091
Db	1139	GAGGAGAAGGGAAGTGAAGCTGTGGGGCAATGTTTTAGAGGCCATACCCATGCTAT	1198
QY	1092	CCACCAAGAGTTAAATTTAATAAACCATTGTTTTTCTGATGATCAGAGCAACATAA	1151
Db	1199	CCCCCGAGGTCAAGTTCAACAACCCCTTGTCTCTTATGATTAAGCAAAATACCAAG	1258
QY	1152	AGCCCATGTTTATGGGTAGGTTGTCAACCCCAACTCAGAA	1192
Db	1259	TCCTCCCTCTTCTATGGGAAAGTGGTGAATCCCAACCCAAA	1299
RESULT 14			
104196			
LOCUS			

DEFINITION Sequence 3 from Patent EP 0137633.
ACCESSION I04196
VERSION I04196.1 GI:591838
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1434)
AUTHORS Parker, M.L. and Kawasaki, G.H.
TITLE Method of expressing alpha-1-antitrypsin in bacteria and its use in therapeutic formulations, and vectors and bacteria for such method and their production
JOURNAL Patent: EP 0137633-A1 3 17-APR-1985;
FEATURES Location/Qualifiers
 source 1..1434
 BASE COUNT 369 a 432 c 340 g 293 t
 ORIGIN
 Query Match 28.1%; Score 429; DB 6; Length 1434;
 Best Local Similarity 60.2%; Pred. No. 2.6e-81;
 Matches 711; Conservative 0; Mismatches 470; Indels 0; Gaps 0;
QY 12 GAAGACCTCAAGGCGACGCGCTCAAAAAACCGACACAGCATCATCCACCATGACGATCAC 71
Db 119 GAGGATCCCGAGGAGATGCTGCCAGAGACAGATACATCCACCATGATCAGGATCAC 178
QY 72 CGGACTTTTAAATAAATTAATCTCCAAATTTAGCCGAATTTGCTTTTCTTTGTATAGACAA 131
Db 179 CCAACTTTCAACAAGATCACCCCAACCTGCTGAGTTGCGCTTTCAGCCTATACCGCCAG 238
QY 132 TTAGCTCATCAAGTAATTTCTACTAAACATTTTTTTTGTAGTCTGTTTCTATTGCCACTGCT 191
Db 239 CTGGCACACCACTGCTCCACAGACCAATATCTTCTCTCCCGAGTGAGCATCGCTACAGCC 298
QY 192 TTCGCACTGTTGAGTTTAAAGTACTAAAGCCGATACCCATGACGAGATTTTAAAGGTTTA 251
Db 299 TTTGCAATGCTCTCCCTGGGACCAAGGCTGACACTCAGATGAATCTCTGGAGGCGTG 358
QY 252 AACTTTAAATTTAGCGAAATCCAGAGCCCAAAATTCAGAGGGTTTTCAGAGAGTTGTTG 311
Db 359 AATTTCAACCTCACGGAGATTCGGGAGGCTCAGATCCATGAAGGCTTCCAGGAACCTCCTC 418
QY 312 AGAACTTTGAATCAACCTGATTTCTCAATTTGAATTAATTAATTAATTAATTAATTTT 371
Db 419 CGTACCTTCAACGACCGACAGCGCTCCAGCTGACCCAGGCAATGGCTGTTCTCCTC 478
QY 372 TCTGAAGGTTTAAATTTGGTTGACAAATTCCTAGAAGAGCTCAAGAAACTATATCATAGT 431
Db 479 AGCGAGGCGCTGAAGCTAGTGATGATTTTGGAGGATGTTAAAGTTGTACCACTCA 538
QY 432 GAGGCTTTTACCGTTAAATTTGGTGATACTGAGGAGCTAAAGCAAAATTAATGATTA 491
Db 539 GAAAGCTTCACTGCTCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTA 598
QY 492 GTTGAAGAGGTCACCGAGGTAAGATCGTTGACCTAGTTAAAGAAATTAAGCTGTA 551
Db 599 GTGGAGAAGGTTACTCAAGGGGAAATTTGGATTGGTCAAGGAGCTTGACAGAGACA 658
QY 552 GTCTTCGCACTAGTTAACTATATTTTTTCAAGGGTAAGTGGGAACTGCTTTCAGG 611
Db 659 GTTTTGTCTCTGGTGAATTTACATCTCTTTAAAGCAAAATGGAGAGACCTTTG 718
QY 612 AAAGATCTGAAGAGGAAGATTTTCATCTTGAATCAAGTTACTGCTCAAAATGTTCA 671
Db 719 AAGGACACCGAGGAAGAGGACTTCCACGTGGACCGTGAACCGGTGAAGGTCCTAT 778
QY 672 ATGAAAGACTGGGTATGTTCAATTTCAACATTCACAAATTAAGTTCTTTGGTCTTA 731
Db 779 ATGAAGCGTTTAGGCATGTTTAAACATCCAGCACTGTAAGAAGCTGTCCAGCTGG 838
QY 732 TTAATGAAGTATTTAGGTAAAGCTACTGCTATTTTTTTTTTACCAGAGCAAGGTA 791

Db 839 CTGATGAATACCTGGGCAATGCCACCGCCATCTTCTCTCTGCTGATGAGGGAACTA 898
Qy 792 CAACATTTAGAGATGAGTTGACTCATCACATTAATAAATTTTATAGAACAGGAT 851
Db 899 CAGCAGCTGGAAATGAATCACTCACCACGATATCATACCAAGTTCCTTGGAAATGAAGAC 958
Qy 852 CGTGTAGCGTTCTCTGCACCTCGCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 911
Db 959 AGAAGTCTGGCAGCTTACATTTACCCAACTGTCCATTACTGGAACCTATGATCTGAAG 1018
Qy 912 TCTGTTTATAGCCAGTTAGTATTAACCAAGTCTTTTCTAACGGTGGCGATTTGAGTGGT 971
Db 1019 AGCGTCTGGTCACTGGGCACTACATAGGTCTTCAGCAATGGGCTGACCTTCCGGG 1078
Qy 972 GTTACTGAAGAGTCTCAATTAATTAAGTAAAGTGTTCACAAAGCGCTTTAACTATT 1031
Db 1079 GTACAGAGGAGGCCCTCAAGCTCTCCAAAGCGCTGCAAGGCTGTGCTGACCATC 1138
Qy 1032 GATGAAGAGGTACCGAGCGCGCGCTATGTTCTGGAAGCTATTCCAATGAGCAT 1091
Db 1139 GACGAGAAAGGAGTGAAGCTGTGGGCGCATGTTTATAGAGGCCATACCCATGTCTATC 1198
Qy 1092 CCACAGAGTAAATTAATAAACCATTCTGTTTCTGATGATCGAGCAACACTAAA 1151
Db 1199 CCCCAGGAGTCAAGTTCACAAACCCCTTGTCTTCTTAATGATGAACAAATACCAAG 1258
Qy 1152 AGCCCATTTTATGGTGAAGTGTCTCAACCCAACTCAGAA 1192
Db 1259 TCTCCCTCTTCATGGGAAAGTGGTGAATCCACCCAAAA 1299

RESULT 15
I04272
LOCUS I04272 1434 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 1 from Patent EP 0139383.
ACCESSION I04272
VERSION I04272.1 GI:591814
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1434)
AUTHORS Russell,P.R.
TITLE Method for expressing foreign genes in schizosaccharomyces pombe and the use in therapeutic formulations of the products, DNA constructs and transformant strains of schizosaccharomyces pombe usable in such method and their preparation
JOURNAL Patent: EP 0139383-A1 1 02-MAY-1985;
FEATURES
source Location/Qualifiers
1..1434 /organism="unknown"
BASE COUNT 369 a 432 c 340 g 293 t
ORIGIN

Query Match 28.1%; Score 429; DB 6; Length 1434;
Best Local Similarity 60.2%; Pred. No. 2.6e-81;
Matches 711; Conservative 0; Mismatches 470; Indels 0; Gaps 0;

Qy 12 GAAGACCTCAAGGCGAGCGCCCTCAAAAAACCGACACAGTCAATCAGCAACAGACCAT 71
Db 119 GAGGATCCCGAGGAGATGCTGCCAGAGACAGATACATCCACCATGATCAGGATCAC 178
Qy 72 CGGACTTTTAAATAAATTAATTAAGCGAAATTTGCTTTTCTTTTGTATAGACAA 131
Db 179 CCAACCTTCAACAAGATCAGCCCAACCTGGCTGAGTTTCGCTTCAGCCTATACGCCAG 238
Qy 132 TTAGCTCATCAAGTAATTTCTACTAACATTTTTTTTAGTCTGTTTCTATTGCCACTGCT 191
Db 239 CTGGCACACCACTCCAAACAGCAGCAACATATCTTCTCTCCCGAGTGAGCATCGCTACAGCC 298
Qy 192 TTTCGGCATGTTGAGTTTAGTACTAAAGCCGATACCCAGATCGAGATTTTAAAGGTTTA 251
Db 299 TTTGCAATGCTCTCCTGGGGACCAAGGCTGACACTCAGATGAATCTCTGGAGGCGCTG 358

Search completed: December 6, 2002, 22:41:39
Job time : 2697 secs